DOCUMENT RESUME

ED 237 582 TM 832 019

AUTHOR Schneider, E. Joseph
TITLE Researcher Questions Use of Standardized Test Results.
INSTITUTION Council for Educational Development and Research, Washington, D.C.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE [78]
GRANT OB-NIE-G-78-0202
NOTE 7p.
PUB TYPE Viewpoints (120) -- Journal Articles (080)
JOURNAL CIT Educational R&D Report; v1 n3 p4-9 1978

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; Educational Change; Educational Innovation; Elementary Secondary Education; Influences; Interviews; Opinion Papers; Program Evaluation; *Standardized Tests; Student Evaluation; *Summative Evaluation; *Testing Problems; Test Results; *Test Use
IDENTIFIERS Cooley (William W); PF Project

ABSTRACT

The results of standardized tests are misused and often do more harm than good, according to William W. Cooley, professor and researcher at the University of Pittsburgh (Pennsylvania). The results seldom assist anyone in improving instruction. Too many out-of-school variables such as social-economic status influence test scores. Achievement is influenced by family background and what the student is exposed to in class; knowing these variables will tell more than tests can. Students tend to learn what they are exposed to; therefore, tests must measure what is being taught in order for them to be useful. Too often this is not the case. The effectiveness of the teaching or curriculum is questioned when, in fact, what is being taught is not being tested. In evaluating programs, schools seldom use tests that measure what the developers claim they are putting across in their materials. Additionally, the tests do not measure the knowledge that a student already possesses or how much time is spent in learning. Summative evaluation isn't answering the questions it is supposed to. All of this is already apparent to practitioners. Given the present state of evaluation, trained classroom observers are much better equipped to explain what works than any available test. (DC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED237582

Researcher Questions Use of Standardized Test Results

E. Joseph Schneider

c.i.	f	d	fd
70–74	1	6	6
65–69	2	5	10
60–64	2	4	8
55–59	3	3	9
50–54	3	2	6
		1	5
			(44)
		0	0
		− 1	− 8
		− 2	− 6
		− 3	− 9
		− 4	− 12
		− 5	− 10
		− 6	− 6
		− 7	− 7
		− 8	− 8
			(− 66)

M = A.M.

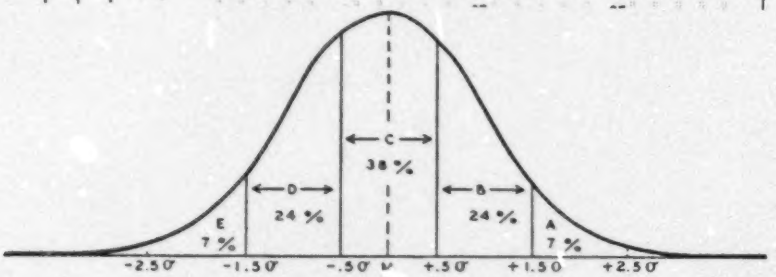
M = Mea
A.M. = Assu
Σfd = Alge
tions about
N = Num
c.i. = Class

M = 42.0

= 42.0

= 42.0

= 39.8



$$n = l. + \left(\frac{\frac{N}{2} - F}{f_m}\right) \times c.i.$$

$l.$ = lower limit of class interval within which median falls

N = one-half of the scores

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☑ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

TM832019

William Cooley

"Actually, the interpretations placed on standardized tests often do more harm than good"

When you come right down to it, it's difficult to talk about standardized testing without getting personal. Often, I've discovered, people's views toward standardized tests depend on what they recall having scored on them.

The people with the scores that ranked them in the upper 99th percentile of their age group, I suspect, tend to be the best supporters of the tests.

Then there are people like me. I remember my standardized test scores. Consequently, long ago I began questioning whether standardized tests can tell me anything worth knowing.

But my convictions were recently shaken. I have a third-grader. She's bright; her mother and I both agree. Thus when her school informed us last winter that all third-grade pupils in the county were about to take a national standardized test, we reassured each other that the test results would be unimportant; we know how smart she is.

Her test scores came in the mail the other day, along with the test developer's completely uncomprehensible explanation about what the scores mean. No matter; I simply looked at the bottom line. My daughter's scores exceeded my expectations.

Her scores, matched against my own experience with such tests, rattled me. If I credit the test with documenting her academic superiority, then I'm forced to admit that my old scores probably labeled me for what I was: an average student.

Confused, I turned to the research community for a sensible answer. While fumbling for a directory of r&d specialists, I spotted a book on my desk, *Evaluation Research in Education: Theory, Principles, and Practice*, by William W. Cooley and Paul R. Lohnes. Of course, Bill Cooley. I knew he's spending some National Institute of Education money looking at instructional procedures that assist in the learning of basic skills. I also recalled that he's been authoring some papers about the evaluation-testing issue.

Cooley is a professor in the School of Education at the University of Pittsburgh and a researcher with the Learning Research and Development Center. The last time I talked to him he had just stepped down as co-director to devote full time to his research program.

So I gave him a call and explained my dilemma. Sure, he said, he would be willing to give me his point of view about testing. But before we arranged a meeting time, Cooley got me to agree to a reading assignment. He wanted me to be familiar with his current work and the fact that his views on testing, while hardly unique, are not shared by all his colleagues.

Test results misused

Cooley's work is more complicated than I had suspected. He and his associates are working to explain why some students within a classroom learn more than other students, why students in some classrooms learn more than students in other classrooms, and why students in some schools learn more than students in other schools. Now those studies get right at the heart of my concerns. Standardized tests, like the one administered to my daughter, often are used to compare her results with those of third-grade students in her own class, other classes, and other schools.

So my first question to Cooley was put rather bluntly: "Do standardized tests prove anything?" I hastened to add that I really wanted to know just what I, as a parent, or an educator, could get from reviewing test results.

Cooley had no difficulty with my question. "Standardized tests have no value at all if what you want to know is which school, for

E. Joseph Schneider is executive director of the Council for Educational Development and Research, Washington, D.C.

example, is doing a 'better' job of teaching kids. In fact, as far as I'm concerned, schools would be better off not giving such tests at all. Actually, the interpretations placed on standardized tests often do more harm than good."

I then asked if there is any justification for schools using national achievement or assessment tests. "Force me into answering your question directly and I'll say there is too little justification for district, state, or national testing or assessment," responded Cooley. "The results can be misused and they seldom assist anyone in improving instruction."

Family background, school learning

"Out-of-school variables such as the students' social-economic status greatly influence scores on such tests"

Getting back to my own situation, Cooley began by explaining that standardized achievement tests are not intelligence tests. Rather, they attempt to record what a student has learned.

"But that's where the problem with the tests begins," he continued. "Let's say, for the sake of discussion, that you and your daughter have equal intelligence. Okay, the difference in your test scores, then, might be attributed to the curriculum you were exposed to. But even that's too simple. We also know it relates to your background. That is, what did you bring with you into the classroom by way of preparation? For instance, out-of-school variables such as the students' social-economic status greatly influence scores on such tests."

And if all that wasn't confusing enough, Cooley added still another variable. It alone, he explained, is sufficient to throw every other variable out of whack: how much time did you spend in the classroom learning the skills that the test measures?

"In other words," Cooley said, "the research community seems to agree that two things influence student achievement: the student's family background (parents' education, home environment, etc.) and what is actually taught in the classroom. Parents and educators alike tend to draw too many conclusions from tests that simply are not justified. Tests cannot tell us as much as simply knowing the child's background and what he or she is exposed to in the classroom."

While Cooley went on talking I allowed my mind to wander a moment. How much time did I spend in class being exposed to what was on my standardized test? It's been too long: I couldn't remember. But I did recall that my daughter's entire school spent a week prior to the test "learning" how to take it. The classes were drilled in skills designed to improve their scores. Once the tests were over, I assumed, the kids went back to learning whatever the school thought was important.

Children tend to learn what they have an opportunity to learn

Turning back to Cooley, I heard him say, "The way to achieve high scores on standardized achievement tests is to bias the instruction in the classroom toward the items being tested and then select kids to take the test who have a high social-economic status." That fit. A good many of the kids with whom I was tested had working-class parents with less than an eighth-grade education. My daughter's classmates, on the other hand, come from a community where the average parent is a professional with college work beyond a bachelor's degree. That background, combined with instruction structured to produce high test results, certainly proved to be a winning formula for success in my daughter's school.

Unwarranted responses to scores

Moving away from my personal concern about test results, Cooley launched a discussion about the overall impact of standardized testing on American education.

The whole point of testing, Cooley reminded me, is to find out whether or not students are learning a set of skills. If the test scores are low, thus indicating to some that students are not learning what we think they should, the logical reaction is to change the curriculum, change the teachers, or write off the students as hopeless learners.

"None of these responses is warranted, based on the tests we are now using," Cooley warned.

Children tend to learn what they have an opportunity to learn, he repeated. By the same token, they tend not to learn things kept from them.

When I observed that such a proclamation was hardly earth-shaking, Cooley readily agreed. "But what's significant is how easily we forget that simple fact when drawing inferences from student test results."

In other words, he explained, achievement test data are worthless to educators and parents alike unless they also know for sure the kids have had an opportunity to learn the things the test is quizzing them about.

"Much too often, educators, school boards, and parents end up questioning the effectiveness of the teaching or curriculum when in fact what is being taught in the classroom isn't being tested. This situation cuts both ways. We don't get a true measure of what the kids have learned. Furthermore, because we often test them on things they never were exposed to, we think they are not learning."

A simple solution, I suggested, would be to "teach to the test."

That's correct, said Cooley, if you want high test scores, rather than an indication of what kids are learning in schools.

A question of what to measure

The whole evaluation-testing picture became even murkier when Cooley went to on explain that we really don't know which educational outcomes to measure.

"I don't want to go around condemning the whole bunch. But they clearly have oversold their product"

"What is it we really want to happen as a consequence of kids and teachers being together in a learning situation?" he asked. "Well, I assume we want the students to improve their performance based on predetermined outcomes. These outcomes generally can be described as increased performance in subsequent schooling or success and satisfaction outside of school."

Well, that sounds logical enough. Logical, yes—but not easy, Cooley continued to explain. Selecting appropriate measures of student outcomes is a problem that continues to plague educational evaluation.

"Because we in education don't know how to measure student outcomes, we generally turn to our handbooks to see what the curriculum developers said were their anticipated outcomes when they produced the classroom materials.

"In other words, we take the developers' stated objectives, measure them, and then describe the extent to which the programs achieved the intended outcomes. When we are all through we feel confident telling ourselves whether or not the program worked, or the kids learned, or the teachers succeeded. But in fact we generally have not answered any of these questions."

The problem, he repeated, is that schools seldom use a test that measures what the program developers claim they are putting across in their materials. And the test results don't tell us anything about what kids brought with them into the classroom to begin with; for example, how much of the information did some of them already have? Also, the test results don't reveal at all how much time a student, or a class, or an entire school district, actually spent mastering the skills in the curriculum.

"Most people, both educators and parents alike, want to know which educational programs are effective," explained Cooley. "And they rely on test results to tell them. But determining the effectiveness of any educational program is difficult and relying merely on a standardized test for the answer is generally misleading.

At this point I was looking for a bottom-line statement. "Are we wasting millions of dollars a year on school testing?" I asked. "And if we are, who's to blame?"

Cooley didn't hesitate with the first question. "Yes, we are

wasting millions on testing. Summative evaluation just isn't working. It's not answering the questions we think it is."

But Cooley held back from a blanket censorship of the huge testing industry in this country. "I don't want to go around condemning the whole bunch. But they clearly have oversold their product."

Skeptical school practitioners

I suggested to Cooley that his views might be hard for the school community to accept. After all, for better or for worse schools are testing students and supposedly making changes in curriculum or instructional approaches as a consequence.

"You might be surprised," said Cooley, "to learn that a good many school practitioners already are aware of these issues. When I discuss the testing controversy with teachers and administrators they say, 'You aren't telling us anything new; we agree with you. We just wish our school board could hear you.' The teachers know, probably better than anyone else, that the students' background and what they spend time studying in the classroom are the major predictors of test scores."

"When I discuss the testing controversy with teachers and administrators they say, 'You aren't telling us anything new; we agree with you' "

Recapping the current situation, Cooley noted that right now we have hundreds of studies using test scores to evaluate educational programs. These studies are going on at all levels: local, state, and national. And a highly skeptical audience of school practitioners is under pressure to make changes as a consequence of the test scores. Many practitioners aren't, of course, because they recognize all the limitations of testing.

"We have to get to the point where the educational decision-makers have confidence in the evaluation-testing results," explained Cooley. "Right now they are skeptical of our work, and they deserve to be."

But Cooley hastened to add that both the decision-makers and the evaluators have to be honest with each other before any meaningful change can occur.

"Too often evaluators—both those employed by a school district and outsiders brought in to conduct a study—simply measure the outcomes specified by the client being served. This practice has to stop. That is, the evaluators must start explaining to the client what can and what cannot honestly be measured. By the same token, the client should be sure that the evaluators are actually measuring the outcomes that the client identifies. In other words, in order to create consensus about what evaluation can accomplish, it's essential to get a good dialogue going between evaluators and clients."

An alternative approach

With that said, Cooley spent a few minutes talking about how his work at the Learning Research and Development Center relates to his overall concern about testing and evaluation.

"As I have stated, I have a real concern that we are not doing a very effective job of measuring what actually occurs in the classroom. Consequently, at the Center we are trying to establish a more valid basis of attributing student growth to particular classroom or school processes. We call this alternative basis or approach, Explanatory Observational Studies.

"We are trying to establish a more valid basis of attributing student growth to particular classroom or school processes"

"We are attempting to understand the major causal networks that operate in a school system where an educational innovation is being introduced. In other words, we are looking at the factors that influence the chances of an innovation having the impact intended. And then, we hope we'll be able to improve our ability to assess the extent of this impact."

If Cooley's inquiry is successful, it may become possible for educators to make programmatic decisions based on evidence that actually establishes the effective dimensions of an educational program. Right now, Cooley believes, that's impossible.

In the meantime, Cooley thinks educators, policy-makers, and parents should remain somewhat skeptical about test results and evaluation studies that make claims for or against the success of an educational innovation.

"Given the present state of evaluation," Cooley concluded, "trained classroom observers are much better equipped to explain what works and what doesn't work than any test we have available."

"Trained classroom observers are much better equipped to explain what works and what doesn't work than any test we have available"

Following my conservation with Cooley I decided to share my daughter's test results with her. I did it with all kinds of caveats about how she shouldn't make too much of her high score. She listened patiently and nodded her head a few times at points in the discussion that seemed to warrant some sign on her part that meant she understood what her father was saying.

What I explained was that her test scores were a whole lot more impressive than mine, but that didn't necessarily mean that one of us was superior to the other. Home factors, I said, make big difference in test results. And so does what was taught to each of us in our respective classrooms.

When I finished, feeling pretty good about my explanation, she simply gave we a hug and said, as only an eight-year-old can, "Don't worry about it, Dad. You are better at some things than I am. I'm just better at taking tests."

Next year I'm going to ask Bill Cooley to explain the test results to her.

For more information please contact the Public Information Office, Learning Research and Development Center, University of Pittsburgh, 3939 O'Hara St., Pittsburgh, PA 15260.

R+D Report p 9